



Opinion Paper

Foundations of Narrative Bibliometrics



Daniel Torres-Salinas^a, Enrique Orduña-Malea^b, Ángel Delgado-Vázquez^c,
Juan Gorraiz^{d,*}, Wenceslao Arroyo-Machado^a

^a Departamento de Información y Comunicación, Universidad de Granada, Spain

^b Departamento de Comunicación Audiovisual, Documentación e Historia del Arte, Universitat Politècnica de València, Spain

^c Biblioteca Universitaria, Universidad Pablo Olavide, Spain

^d Vienna University Library, University of Vienna, Austria

ARTICLE INFO

Keywords:

Responsible metrics

DORA

CoARA

Evaluative bibliometrics

Narrative curriculum

Narrative Bibliometrics

Databases

Bibliometric indicators

Narratives

ABSTRACT:

The document 'Foundations of Narrative Bibliometrics' provides an analysis of the evolution of scientific assessment, highlighting the influence of manifestos such as DORA and CoARA in shaping ethical and responsible practices in academia, as well as their assimilation by Spanish scientific policies. It connects this context with the contributions of evaluative bibliometrics, emphasising the transition towards a more integrative approach that advocates for a balance between quantitative and qualitative methods in research evaluation. Furthermore, it underscores how the Narrative Curriculum has emerged as one of the fundamental tools in new evaluation processes, as it allows for the description of the complexity and context of academic achievements. Narrative Bibliometrics is proposed, defined as the use of bibliometric indicators to generate stories that support the defence and exposition of a scientific curriculum and/or its individual contributions within the framework of a scientific evaluation process, which demands narratives. To introduce the reader, it presents, in a non-exhaustive manner, sources, indicators, and practical cases for effectively applying Narrative Bibliometrics in various scientific evaluation contexts. Hence, this document contributes to the responsible use of bibliometric indicators, serving as a tool for evaluators and researchers.

1. The new evaluative framework

1.1. The winds of change

In the realm of scientific evaluation, a new model of assessment has been developed over recent years, articulated in three clearly defined phases. The first phase is characterised by the publication of a series of manifestos and declarations, culminating in the consolidation of a movement towards the responsible use of metrics (Pérez Esparrells et al., 2022). In the second stage, the assimilation of the manifestos is observed, along with a growing commitment by some institutions towards the development of more sophisticated indicators and ethical evaluative practices more adapted to the specific context of evaluations. The third phase involves the institutionalisation of these principles through the creation of alliances such as CoARA (Coalition for Advancing Research Assessment),¹ representing the formalisation of the Open Science movement's aspirations and values into robust organisational structures aimed at

* Corresponding author.

E-mail address: juan.gorraiz@univie.ac.at (J. Gorraiz).

¹ <https://coara.eu>

promoting systemic and global changes beyond mere publication and data openness (Abadal, 2021), but at the very core off the scientific evaluation process; changes fundamentally guided by the use of qualitative judgements and peer review, to the detriment of bibliometric indicator usage (Torres-Salinas et al., 2023b).

The current moment is marked by the integration of these principles into evaluative practice. In Spain, the effective implementation of these principles by ANECA (National Agency for Quality Assessment and Accreditation),² as detailed in the recent Resolution of December 5, 2023, from the National Commission for the Evaluation of Research Activity, publishing the criteria for the evaluation of research activity,³ constitutes a significant case study due to the implications ANECA has as the agency responsible for accreditations and tenure in Spain and, therefore, as an inspiration for other national scientific policies.

Bibliometrics, as a field of study, has significantly evolved in the last part of the 20th century, being consolidated as an applied discipline initially based on the counting of citations and publications. Throughout the 20th century, the process of retrieving bibliographic citations in scientific works was monopolised by a single company,⁴ which commercialised the so-called Citation Indexes (Science Citation Index and Social Science Citation Index), later known as Web of Science (WoS), as well as the well-known Journal Citation Reports (JCR), which included the now repudiated and criticised Journal Impact Factor (JIF) as its flagship indicator.

Bibliometric indicators were adopted *en masse* at the expense of an exclusive dependence on obtaining them from a single data source (i.e., WoS), despite its obvious limitations in bibliographic coverage (Mongeon & Paul-Hus, 2016). The main consequence has been a predominant and decisive use of the Journal Impact Factor (JIF) in evaluating the scientific impact of articles and researchers by evaluation agencies and scientific institutions, despite the recommendations of its own creator (Garfield, 2002, 2006), as well as hundreds of works and editorials that denounced the limitations linked to its misuse (Larivière & Sugimoto, 2019; Bornmann & Williams, 2017). This discourse also proposed improvements in its calculation (Thelwall & Fairclough, 2015) or outright suggested other better-designed indicators that could replace it (Leydesdorff, 2012; Moed, 2010).

It is clear that practitioner scientometricians have been among its most vocal critics, but its large-scale implementation was not their business. However, they have clearly failed in bringing their concerns and criticism to bear outside the purely academic exchange into the real world of research evaluation beyond their own involvement. Despite this, the JIF consolidated its reign based on its ease of understanding, apparent low cost to the end user, and speed of calculation, promoting an extensive and global use of metrics. Nonetheless, from the year 2004 onwards, the situation began to change.

On one hand, a greater number of information sources were introduced to the market,⁵ allowing for a diversification in citation acquisition, which in principle, mimicked the existing model and replicated the main indicators with subtle differences. Moreover, the digital era and Web 2.0 contributed to the emergence of new digital metrics (Bollen et al., 2005), expanding the quantitative horizons towards so-called webometrics (Almind & Ingwersen, 1997; Thelwall, 2008), and ultimately generating a complete and infinite family of alternative indicators (altmetrics) (Priem & Hemminger, 2010; Torres-Salinas et al., 2013). This transformation led to a significant increase in the supply of sources, indicators, methodologies, and techniques, also favouring quantitative evaluation at the contribution level, thereby reducing the dependence on journal-based metrics or other publication aggregates.

This context led to reflection processes that resulted in the publication of various manifestos and declarations. These manifestos primarily aimed to criticise certain bibliometric practices and indicators, especially the JIF and the H-index. They also encouraged the adoption of good practices in the use of bibliometric indicators, promoting their adaptation to the new realities of scientific research and the incorporation of new metrics, more in line with the new platforms for the publication and dissemination of scientific knowledge. Among the most influential and decisive documents are the San Francisco Declaration on Research Assessment (DORA)⁶ and the Leiden Manifesto⁷ (Hicks et al., 2015).

- *DORA (San Francisco Declaration on Research Assessment)* (2012): This Declaration advocates, in the metric part, for the elimination of the emphasis on the JIF as the main reference and focuses on evaluation based on merit and the actual impact of research. DORA is one of the first Declarations to address the need to change the way research is evaluated at a global scale.
- *Leiden Manifesto* (2015): This manifesto (Hicks et al., 2015) focuses on providing principles for research evaluation and decision-making based on bibliometric metrics. It advocates for transparency in the selection of indicators, the importance of diversity in evaluation methods, and the need to consider the disciplinary context in research evaluation.

While Leiden and DORA have been the most relevant activities, they have not been the only ones, being complemented by other sector-specific manifestos, among which the Altmetrics Manifesto⁸ and the Metric Tide report (Wilsdon et al., 2015) particularly stand out in our field.

- *Altmetrics Manifesto* (2010): The Altmetrics Manifesto (Priem et al., 2011) highlights the importance of alternative metrics (altmetrics), focused on expanding the impact dimensions of scientific publications through the quantification of their dissemination

² <https://www.aneca.es>

³ <https://www.boe.es/boe/dias/2023/12/16/pdfs/BOE-A-2023-25537.pdf>

⁴ The Institute for Scientific Information was later transferred to Thomson Reuters and is currently owned by Clarivate Analytics.

⁵ In 2004, Google Scholar (Alphabet) and Scopus (Elsevier) were launched.

⁶ <https://sfidora.org/>

⁷ <http://www.leidenmanifesto.org/>

⁸ <https://altmetrics.org/manifesto/>

and consumption on websites and social media platforms, as well as their mention in media and non-scientific publications (reports, guides, grey literature), through unconventional indicators beyond traditional citation metrics.

- *The Metric Tide* (2015): An independent report by the UK Metrics Task Force on the ethical and social challenges related to research metrics of significant influence among bibliometricians. It emphasises the importance of more responsible and ethical research evaluation, highlighting the need to focus on quality and diversity of outcomes, promoting transparency and accountability in the use of metrics, and advocating for a multidimensional evaluation that includes both qualitative and quantitative indicators. Additionally, the report warns about the unwanted effects of an excessive dependence on metrics and hypercompetition in scientific research.

Table 1 summarises in 10 points the key aspects of the different manifestos which, to some extent, determine the new contexts.

1.2. Old practices

These advancements or consensuses within the scientific community did not emerge spontaneously. Much of these practices, as we argue in an upcoming monographic work (Torres-Salinas, 2024), are indebted to and draw from Evaluative Bibliometrics, a specialty with a defined and indisputable origin in the National Science Foundation report titled “Evaluative Bibliometrics” (Narin, 1976). Subsequently, Evaluative Bibliometrics was systematised during the ‘80s at the CWTS (Centre for Science and Technology Studies) (Petersohn & Heinze, 2018), whose members defined its principles in multiple works, and from which we can find recent syntheses (Daraio & Glänzel, 2020; Moed, 2017a; Van Raan, 2019).

It should be noted that Evaluative Bibliometrics emerges as an instrumental discipline to support Peer Review (Table 2), not as a substitute for human judgment, and is focused on making decisions responsibly. Obviously, if there were no decisions regarding indicators, we would be talking about Descriptive Bibliometrics. From its origins, therefore, Evaluative Bibliometrics has promoted a fair, rational, and limited use of bibliometric indicators and has also opposed the so-called ‘quick and dirty’ metrics, characteristic of Desktop Bibliometrics (Moed, 2020) and salon metrics (Aguillo, 2022).

In essence, the principles advocated by Evaluative Bibliometrics (EB) since its origins in the ‘80s are five:

- **Principle of support for decision-making:** bibliometric indicators are a tool to assist in decision-making in scientific policies. The essential goal is to provide objective indicators for effective evaluation.
- **Principle of collaboration with experts:** Bibliometricians work at the service of experts from different disciplines, who use the indicators as a complement to their qualitative judgements. Indicators are subordinate and complementary to Peer Review, not a replacement for it.
- **Principle of respect for contexts:** the evaluation process and the indicators must respect the context of the evaluated agents, considering the evaluation framework, career, discipline, language, or type of publication.
- **Principle of metric multidimensionality:** there are indicators capable of measuring different aspects of impact, such as scientific-academic, educational, economic-technological, or social and cultural.
- **Principle of verifiability and data openness:** transparency, accessibility, and verification of data are necessary, ensuring their reliability, so that they provide guarantees to both evaluators and those being evaluated.

1.3. New policies

In the previous points, we have described the general context that allows us to understand the emergence of CoARA - Coalition for Advancing Research Assessment in December 2022, an initiative strongly influenced by DORA, seeking to reform the research evaluation system. CoARA is based on an agreement, the Agreement on Reforming Research Assessment, which establishes common commitments respecting the autonomy of organisations. The coalition was driven by the European Science Foundation – Science Connect (ESF-SC)⁹ and the European Universities Association,¹⁰ with the support of the European Commission, and currently brings together more than 650 organisations, including universities, R&D centres, and evaluation agencies.

Among the foundations of CoARA, it is explicitly mentioned that emphasis will be placed on qualitative aspects and the diversity of outcomes and that, therefore, research evaluation should be based primarily on qualitative judgements, with peer review as a central element, supported by a responsible use of quantitative indicators. Furthermore, CoARA advocates for a more balanced and qualitative approach in evaluation, recognising the diversity of outcomes and practices in research. This implies a change from the current dependence on quantitative indicators, seeking a more holistic and representative evaluation.

Therefore, we observe that both Evaluative Bibliometrics and CoARA are based on the same principles: evaluations should be based on expert qualitative judgements, which should be complemented with bibliometric indicators. A critical aspect to highlight from this expert-based perspective is the uncritical adoption of peer review by CoARA. The limitations of this process are not detailed, suggesting an overly idealized view (Torres-Salinas et al., 2023b). Scientific literature extensively documents the biases of peer review. For example, Lee et al. (2013) identified several significant biases, including confirmation, conservatism, affiliation, gender, nationality and language, prestige, and bias against interdisciplinary research. In this context, it would be prudent for CoARA to include a detailed

⁹ <https://www.esf.org/>

¹⁰ <https://eua.eu/>

Table 1
Summary of the main points from the metric manifestos.

	• Strongly criticises the use of the JIF of journals for funding and academic promotion decisions. • Encourages institutions and funding agencies to consider a wide range of metrics and indicators, including qualitative and quantitative metrics that reflect the actual impact.
	• Proposes to recognise and value the diversity of contributions in research, including data, software, methods, and other types of intellectual output that are fundamental to the advancement of science. • Promotes transparency in the publication and evaluation of research, advocating for open access and data sharing. • Emphasises the need to adapt evaluation metrics to the specific goals and contexts of each research, combining qualitative and quantitative methods. • This allows for a more complete and fair appreciation of research performance and impact in its various aspects: academic, social, economic, cultural. • Highlights the importance of transparency in evaluation methods and the accountability of evaluators. • Furthermore, it advocates for the recognition of a variety of scientific contributions (bibliodiversity), beyond publications in high-impact journals.
	• Seeks to broaden the understanding of academic impact beyond citations, including mentions on social networks, downloads, bookmarks, blog discussions, and media coverage, etc. • Altmetrics offer the possibility of dynamic evaluation, providing real-time data on how research outcomes are discussed and used. • Allows for valuing a wider range of academic contributions, such as research data, software, and contributions to online communities. • Altmetrics advocates for the use of open data to facilitate the verification and replication of its metrics, promoting transparency in research evaluation.
	• Advocates for an ethical and responsible research evaluation, avoiding excessive simplification through metrics. • Highlights the importance of considering both the quality and diversity of research outcomes. • Promotes transparency and accountability in the use of metrics, warning about the risks of excessive dependence on them. • Recommends including stakeholders in decision-making on metrics and the continuous review of evaluative approaches.

Table 2
Henk Moed's vision of scientific evaluation in 2007 from CWTs.

<i>"Outcomes of citation analysis must be valued in terms of a qualitative, evaluative framework that takes into account the substantive content of the works under evaluation: this can be done by peers only. The future of research evaluation rests with an intelligent combination of advanced metrics and transparent peer review. It argues that metrics, especially a sophisticated type of citation analysis, may provide tools to keep the peer-review process honest and transparent. Both metrics and peer review have their strengths and limits. A real challenge is to combine the two methodologies in such a way that the strength of the first compensates for the limitations of the second, and vice versa"</i>	Moed, H. F. (2007). The future of research evaluation rests with an intelligent combination of advanced metrics and transparent peer review. <i>Science and Public Policy</i>, 34(8), 575–583. Henk Moed from the CWTs was one of the pioneers and popularisers of Evaluative Bibliometrics through an extensive bibliography spanning more than three decades.
---	--

section on the limitations and inherent biases of peer review in its guidelines.

Despite this, this aspect has gone unnoticed and it is common to find that DORA and CoARA are signed together, an aspect that institutions increasingly communicate publicly on social media platforms as part of their external communication policies (Orduna-Malea & Bautista-Puig, 2024). The adoption of these agreements has intensified, with the signing of both agreements by the European Commission on November 8, 2022, as announced by the Directorate-General for Research and Innovation.¹¹

As a direct consequence, numerous institutions have followed this trend. In the Spanish case, the support for CoARA is beyond doubt, being the country with the most adhesions to the principles, including, among others, significant bodies and institutions such as CRUE, CSIC,¹² and, of course, ANECA, which signed both agreements on April 3, 2023. The immediate consequence is a global redesign of all evaluative practices in its different competencies (accreditations, tenure, etc.). Without mincing words, this new change of direction in Spanish scientific policy was announced on ANECA's website (Table 3).

2. The narrative CV

At this pivotal moment, the question arises about what tools will be available for evaluators/evaluated. One of the significant changes brought by these new evaluation systems is the implementation of the Narrative CV, a philosophy primarily supported by DORA and already beginning to be adopted by many institutions, albeit in a quite heterogeneous manner. An example is the Spanish National Research Agency, which requests Narrative CVs from Principal Investigators (PI) in their research project evaluation

¹¹ In the EU press release on the signing of CoARA and DORA, there are two interesting points. On one hand, the statements by Mariya Gabriel, Commissioner for Innovation, Research, Culture, Education and Youth, emphasize that these signings are "commitments from the European Commission to pave the way towards a reform of research evaluation practices." This means we are at the beginning of a reform, not an end. On the other hand, it is included as an action in the Policy Agenda of the European Research Area (ERA) for 2022-2024, which includes an action to advance the reform of the research evaluation system, researchers, and institutions.

¹² CoARA approves Crue's reform proposal: ANECA and CSIC for Spain <https://www.crue.org/2023/07/coara-aprueba-la-propuesta-de-reforma-de-crue-aneca-y-csic-para-espana/>

Table 3ANECA's statement following the signing of DORA and CoARA in April 2023.¹³¹

"ANECA adheres to DORA because it shares the necessity to address the evaluation of the quality of scientific articles—not just the journals in which they are published—, to consider the value and impact of all research outputs (including data and software), and to consider the social impact of research from a broader perspective (including qualitative indicators, such as influence on policies and/or scientific practices)."

"Likewise, ANECA adheres to CoARA, committed to recognizing the diversity of research practices and activities in evaluation processes, which also maximize their quality and impact, achieving a more efficient and inclusive system. To this end, it commits to progressively moving towards qualitative evaluation models, based on peer review, and with the support and responsible use of quantitative indicators."

processes, as part of the Abbreviated Curriculum Vitae, the compulsory academic résumé in funded national research calls.

This approach aims to move beyond extensive listings of publications and merits. According to DORA's website, narrative curricula promote a "quality over quantity" mindset in the evaluation of academic careers, reducing the emphasis on journal-based indicators and adapting to non-linear research trajectories. Thus, narrative CVs aim to minimize the role of journal prestige in the evaluation of candidate profiles (Hazlett, 2022). The underlying idea is to allow individuals to more effectively demonstrate their contributions to research, teams, and society at large (Byrne, 2023). Currently, the implementation of these curricula varies and is adapted to each organization, resulting in multiple versions. However, the path is not straightforward, and the implementation of narrative CVs is clearly fraught with numerous challenges, including lack of explicit evaluation criteria, inadequate guidance for researchers, and the difficulty in establishing effective narrative assessment systems (Cobey, 2024).

One of the most popular Narrative CVs is known as the Royal Society's Résumé for Researchers (R4R) (UKRI, 2023), a flexible narrative CV template designed to present a wider range of skills and experiences inspired by the "Résumé for Researchers"¹⁴ format, introduced by the Royal Society of London in 2019. R4RI¹⁵ combines elements of a CV and a cover letter, providing space to explain the context of academic achievements, allowing research staff to present relevant skills and experiences including mentoring and leadership, committee membership, outreach activities, and other contributions that do not always result in publications.

Another example of this new format can be found in the Swiss National Science Foundation,¹⁶ initially piloted in 2020 for research in medicine and biology, the curriculum includes eight sections and asks research staff to present four 200-word narratives describing their most significant contributions to science (Singh Chawla, 2022). Thus, it is suggested that these new curriculum formats could promote greater diversity and inclusion in science (Woolston, 2022) since narratives would allow scientists to better explain their achievements and contributions to science, especially when they do not fit into traditional résumé categories.

Therefore, the adaptation and implementation of the narrative CVs proposed by DORA fall to the institutions, although all follow a common line. It is relevant to highlight that the global adoption of these new curriculum formats presents various challenges. One of them is the lack of homogeneity, an aspect against which DORA suggests the need for standardisation. Furthermore, it points out the importance of training both users in their preparation and reviewers in their interpretation. Hence, it proposes developing a common and flexible definition of the narrative CV that is adaptable to different organizational contexts (Fritch et al., 2021). For professionals dedicated to evaluation and to the creation of tools that facilitate this standardisation, we face what we call the "metric challenge." This challenge involves overcoming the dependence on the JIF (not necessarily its complete exclusion) and introducing metrics that strictly adhere to the framework and rules established in manifestos and declarations. Following DORA's guidelines, organizations may employ quantitative indicators, including bibliometric indicators, as long as they comply with the conditions indicated in Table 4.

This provides institutions and applicants with the freedom and opportunity to complement the defense of their merits in a robust and reasoned manner, using metrics and indicators subordinate to other types of narratives. Precisely, it is in this last challenge that Narrative Bibliometrics comes into play, which we will define next as a framework for the use of quantitative indicators conveniently described as a complement to other narratives.

Contrary to what is sometimes maintained, the narrative CV is not entirely foreign to the Spanish framework. Specifically, in evaluative processes such as tenure evaluations and accreditations, free text spaces are used for applicants to develop the narrative of their research careers (as in the PEP program¹⁷ and the previous ACADEMIA format), their research merits or any other type (the current ACADEMIA), or about specific merits (Sexenios¹⁸ and ACADEMIA). In the same vein, applicants for projects from the Spanish National Plan have traditionally been able to narrate the main merits of the research teams, all of which can include the use of metrics. Therefore, we are moving in familiar territory. This manuscript aims to identify, describe and systematize some common practices that have already been taking place in Spanish academia.

¹⁴ <https://royalsociety.org/topics-policy/projects/research-culture/tools-for-support/resume-for-researchers/>

¹⁵ The R4RI template includes a total of four modules: 1) Contributions to the generation of new ideas, tools, methodologies, or knowledge, 2) Development of others and maintenance of effective working relationships, 3) Contributions to the broader research and innovation community, and 4) Contributions to wider audiences and users of research/innovation and towards a broader social benefit. Applicants can provide additional details relevant to their application in a separate section. The complete template of this CV can be consulted at: <https://www.ukri.org/publications/resume-for-research-and-innovation-r4ri-template>

¹⁶ <https://www.snf.ch/en>

¹⁷ The PEP is a program by ANECA aimed at evaluating the teaching and research performance of university faculty members in Spain.

¹⁸ A sexenio is a six-year period used in Spain to assess the research activity of university professors for rewards and promotion.

Table 4
Conditions for the Use of Indicators and Metrics According to DORA.

1	2	3
Minimize the role of the prestige of scientific journals	Privilege the quality and nature of achievements over quantification	Use metrics responsibly

3. Narrative Bibliometrics

Given its conceptual simplicity, Narrative Bibliometrics can be defined as "the use of bibliometric indicators to generate stories and narratives that allow for the defense and exposition of a scientific curriculum and/or its individual contributions within the framework of a scientific evaluation process" (Torres-Salinas, 2023a). Thus, we are talking again about a type of Bibliometrics that is instrumental in nature, essentially acting as a support or aid for the generation of specific quantitative narrative aimed at a detailed and responsible exposition of evidence and indications related to the impact, attention, dissemination, and influence of any type of scientific results. Due to its inherent characteristics, Narrative Bibliometrics must necessarily be framed within the theoretical and practical corpus of Evaluative Bibliometrics, which implies adherence to and respect for the five fundamental principles previously exposed (Section 1.2). This means that the information provided by bibliometric narratives must facilitate decision-making in scientific policies, such as hiring, promotions, and salary supplements. Moreover, these narratives are subordinate to expert judgment and serve as support for other forms of narrative (Fig. 1).

From a methodological perspective, bibliometric narratives must consider the diverse contexts of both the evaluative framework and the subject evaluated. It is essential that they have a multidimensional character, avoiding an exclusive focus on journal indicators or citations. Lastly, the information provided must be transparent and verifiable, thus ensuring its reliability and usefulness in the evaluation process. Therefore, Narrative Bibliometrics is a younger sibling of Evaluative Bibliometrics, whose main goal is to offer practical solutions in new evaluative procedures.

What exactly does Narrative Bibliometrics entail? It can be described as the 'Bibliometrics of common sense', where raw data is given meaning in a logical manner. That is, it is a narrative that encourages researchers to analyze and contextualize each indicator, avoiding their isolated and disconnected presentation. This moves away from what has traditionally been known in Evaluative Bibliometrics as 'bean counting', a term referring to the decontextualized application of indicators (Rafols & Stirling, 2021). While descriptive Bibliometrics has historically been a data-driven discipline, our proposal moves towards an understanding of bibliometric data, recognizing that numbers alone cannot fully capture the value and importance of scientific research.

In a sense, by promoting a narrative of what happens behind the data, Narrative Bibliometrics avoids the simplification involved in merely presenting a bibliometric indicator. Moreover, it can help to prevent problematic behaviors in academia, such as the manipulation of indicators or high rates of self-citation.

Within the framework of Narrative Bibliometrics, it is crucial to establish principles that guarantee not only the quality and relevance of the analysis but also its replicability and adaptability to different academic contexts. These narratives must go beyond mere data presentation, incorporating elements that ensure a comprehensive understanding of the scientific contribution. For this purpose, it is essential to consider key aspects such as the clarity of sources, the uniqueness of contributions, adaptation to the specific field of study, appropriate contextualization, and avoiding direct comparisons with other colleagues. These recommendations are developed in detail in Table 5.

4. The narrative dimensions

There are various ways to construct and understand narratives derived from bibliometric indicators, many of which may overlap and must be combined appropriately to create an effective discourse and defense. Below, we explore five key examples of these narratives, each providing a unique perspective. These examples range from the Narrative of Position, which contextualizes contributions within a broader framework, to the Narrative of Open Science, highlighting proactivity in the dissemination of scientific work. These perspectives enrich bibliometric data and offer a more comprehensive view of the impact of scientific work. It is important to mention that the narratives presented focus on individual contributions. Additionally, it should be noted that curricula vitae (CVs) often have a character limit, implying the need for effectively synthesizing the message for the evaluating personnel.

- **Narrative of position:** Provides an essential comparative perspective to position a scientific contribution relative to others. This narrative relies on the use of normalized indicators, such as the Category Normalized Citation Impact (CNCI) or the Field Weighted Citation Impact (FWCI), which evaluate the impact of a publication compared to other works in the same field, year, and document type. These indicators are particularly valuable as they offer an analysis adjusted to the specificities of each discipline (Torres-Salinas et al., 2018). Furthermore, the utility of these indicators extends when bibliographic databases provide the position of an item in a collection, considering quartiles and percentiles, translating into a significant measure to determine the relevance of a publication within its specific field. This narrative approach is especially valuable for highlighting the position of research through various indicators, whether citation, usage, among others, thereby underlining its importance and relative contribution in the

¹³ [https://www.aneca.es/-/aneca-se-adhiere-a-la-san-francisco-declaration-on-research-assessment-dora-y-a-la-coalition-for-advancing-research-assessment-coara-](https://www.aneca.es/-/aneca-se-adhiere-a-la-san-francisco-declaration-on-research-assessment-dora-y-a-la-coalition-for-advancing-research-assessment-coara)

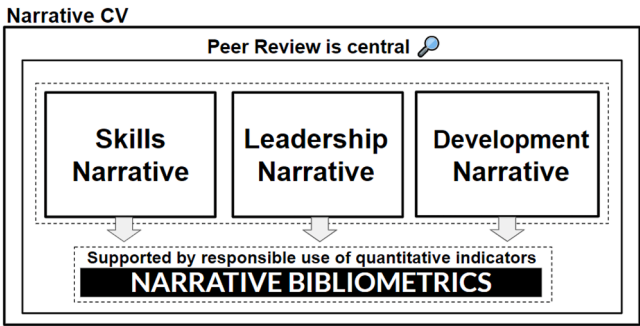


Fig. 1. The integration of Narrative Bibliometrics in the framework of narrative CV established by DORA.

Table 5
General considerations for conducting Bibliometric Narratives.

Replicability	Uniqueness	Adaptability	Comparability	Contextuality
The bibliometric narrative requires a clear and precise description of methodologies and sources. This allows other researchers to replicate the studies and reach similar conclusions.	The narrative should reflect the individual contribution or the specific impact that one wishes to highlight uniquely. Despite similar contributions, the specificity of each contribution should be emphasised. Therefore, the same contribution could be defended in different ways.	The narrative must adapt to the particularities of each field, considering the differences in publication, citation, and collaboration practices that vary between disciplines. Thus, global and uniform indicators and narratives do not apply to all areas.	Narratives should focus on highlighting individual achievements and contributions without making direct comparisons with other colleagues. This promotes an evaluation based on one's own merits and avoids creating an environment of undue competition.	The narrative must place the data and achievements in a broader context. This includes explaining the relevance of the work within the field of study, its impact on the scientific community and society, and how it relates to current research trends.

- scientific field. Metrics referring to the average number of citations available in products like Essential Science Indicators, Hot Papers, and Highly Cited Papers could also be included in this category.
- ✓ *Despite its recent publication, our proposal is receiving significant attention from the scientific community, and the study indexed in Scopus under the thematic category of "Environmental Sciences" has 40 citations, thus significantly exceeding the average of 10 citations typically expected for publications in this category in 2021. Therefore, it quadruples the citation expectations in its field. Applying the Field-Weighted Citation Impact (FWCI) indicator, which contrasts received citations with those expected, results in an FWCI of 4, thus evidencing the exceptional impact of the work in the area of Environmental Sciences.*
 - **Narrative of context:** Emphasizes the relevance of the environment and place in which the publications have been cited, offering an enriching perspective on the application and practical utility of scientific works. By conducting a detailed analysis of the locations and modes in which works are mentioned, a detailed narrative about their theoretical and practical impact can be developed. This approach allows discerning, for example, if a study is having a significant influence on the development of theoretical debates or the evolution of methodologies through the classification of citations according to article sections (e.g., introduction, methodology, discussion). Such distinction is of utmost importance for evaluating personnel, as it provides a deeper and more nuanced understanding of the scope and relevance of research. Contextual metrics, integrated into platforms like Web of Science or the *_scite*¹⁹ database, can be effectively employed for a more comprehensive assessment of academic impact (Orduña-Malea, 2022).
 - ✓ *Our work has received 30 citations. Consulting *_scite*, we find that 5 citations are in the introduction, 5 in methodology, and 20 in the discussion. This indicates that the work has generated some interest, being used as a basis for contrasting results in other publications. For example, Braun (2023) states that 'our results corroborate the values reached for gene sequencing in diverse populations by Torres-Salinas, which means that their method is valid for comparative genomic studies and opens the doors to multiple applications in personalized medicine and the analysis of genetic patterns at a population level.*
 - **Narrative of agents:** Focuses on identifying, describing, and characterizing the different actors, both individual (authors) and collective (groups) and institutional (organizations), that use and employ the results of our work. Providing a detailed analysis of who is applying the knowledge derived from our research gives a clear and objective view of its real impact. This perspective is fundamental to understanding the nature of the impact generated, whether educational, social, economic, or otherwise. Tools like Overton,²⁰ for example, offer the ability to track the use of a publication in globally relevant policy reports, such as situation reports

¹⁹ <https://scite.ai/>
²⁰ <https://www.overton.io/>

from the World Health Organization (WHO) or strategic documents from the European Union. This narrative, therefore, not only allows an understanding of the influence of our publications in various contexts outside the scientific ecosystem.

- ✓ Our publication "The spread of true and false news online" has had a wide dissemination in different impacts to the scientific community as it is referenced in 26 policy reports such as the "OECD Public Governance Policy Papers" (OECD), the "Key social media risks to democracy" (Publications of the European Union), or the "Online Safety Bill: supporting documents" (gov.uk). Therefore, our recommendations on the evaluation of fake news are being applied for the design of public policies and the creation of guides. The indicated information can be verified in the Overton database. Moreover, it has not only had a political influence but also transferred knowledge to the technological field as demonstrated by the four mentions from patents registered in the United States Patent and Trademark Office (US-11636679-B2, Grant US-11494446-B2, Grant US-11176380-B2, Grant US-10691951-B1), information that can be quickly verified through the Dimensions database.
- **Narrative of audience:** Focuses on the potential reach (audience) of findings across different audiences. Unlike the narrative of agents, which makes use of and mentions our publications in other texts, the audience narrative specifically focuses on the reception of our work and its filtering in various ways. It is crucial not only to quantify but also to qualify the recipients of our research. In the context of the press, for instance, it is important to consider both the media outlets involved and their audiences, considering their reach (local, national, international). On digital platforms like X (aka Twitter), not only the volume of mentions is relevant but also the potential reach of these, considering the number of followers of those who interact with the content. The detailed identification and description of the audience profile, whether scientific or media, adds a crucial dimension to the impact analysis. This approach has been discussed by Arroyo-Machado and Torres-Salinas in 2023, highlighting its importance on platforms like X or Wikipedia.
- ✓ The results of our study on adolescent mobile phone use were disseminated on Twitter to inform various associations, professionals, and colleagues. The work was retweeted 65 times, and according to Twitter, the total tweet impressions were 10,100. It's worth mentioning that 40 % of the interactions came from institutional accounts such as @JovenesSaludables, @PrevencionAdol; some scientific societies also echoed it, like @SocPediatria, @CienciaAdolesc, and @PsicoEducaJoven. In the comments from various institutional accounts (for example, @InstFamiliaAdol, @EducaSalud), they mention that the protocols on youth well-being and digital risk prevention are reinforced by our findings. Furthermore, as a result of this dissemination, we were contacted by the EducaDigital Association to offer a virtual course on effective strategies for responsible mobile phone use by adolescents.
- **Narrative of Open Science:** This narrative highlights the research personnel's proactivity in disseminating their work on open platforms. It primarily focuses on indicating whether, on the one hand, publications have been deposited in open access in institutional or thematic repositories (for example, the University of Granada has DIGIBUG, and at the European level, we recommend Zenodo). These narratives can be accompanied, where appropriate, by the repositories' usage data. On the other hand, it focuses on indicating whether the preprint version was deposited and disseminated prior to publication, or if the data derived from our publications (datasets, software) have been deposited in open access. Beyond mere deposit, it can be indicated if the description and deposit of these data have been carried out according to certain standards (for example, the FAIR principles²¹). In addition to the mentioned usage indicators, indicators on file size, shared volume, etc., can be offered. In all cases, it is advisable to accompany the records with persistent identifiers (DOI, handle, etc.) for correct identification.
- ✓ Our recent article on learning algorithms has been disseminated in pre-print format in the Digibug (handle 10.5430/DGB.2023.056AB) and Zenodo (DOI 10.5281/zenodo.789123) repositories. Additionally, the study's supplementary materials are available on FigShare (DOI 10.2132.3232/fig2224). Among these, Supplementary Document 1 (DOI:10.1234/SCI2S.2023.001), offering detailed precision results for each classification method, and Supplementary Document 2 (DOI:10.1234/SCI2S.2023.002), with the Wilcoxon test results, stand out. These documents, with more than 5400 combined downloads (3723 from the United States, 1200 from the United Kingdom), have generated significant interest in the academic field.
- ✓ The datasets from our study on machine learning algorithms, hosted on DataSphereX under the identifier IX:10.DSX/2023/DATA01, comprise approximately 500 GB of information distributed over more than 2 million rows and 100,000 columns. With over 3000 downloads, these data have been fundamental in cross-disciplinary research, complying with the FAIR standards for scientific data management and use. For example, in educational psychology, they have been reused in the study "Impact of Machine Learning on Educational Assessment" (DOI:10.5555/edu-tech.2023.004), analyzing 200,000 rows to discover learning patterns. In communication sciences, they enriched the analysis in "Social Networks and Informational Behaviour" (DOI:10.5555/com-sci.2023.010), and in public health, they contributed to "Big Data Analysis for Disease Prevention" (DOI:10.5555/pub-health.2023.021). This interdisciplinary reuse demonstrates the applicability and relevance of the data, highlighting the importance of sharing open and accessible information to foster scientific collaboration and knowledge advancement.

5. Methodological framework

Narrative Bibliometrics adopts a more holistic and contextual approach to evaluating scientific research. This method transcends the use of a limited set of indicators, focusing on detailed argumentation that highlights multiple dimensions of research. Indicators act as tools within a broad interpretative framework that seeks to comprehensively understand the relevance and impact of research. By incorporating the stories and contexts underlying the data, Narrative Bibliometrics promotes a deeper and more nuanced

²¹ There are tools that automatically validate the compliance of a digital object (dataset) with the FAIR criteria on a pilot basis, such as F-UJI (<https://www.f-uji.net>).

understanding of academic work, encouraging more complete, thorough, and granular evaluations. It involves being able to identify activity and impact in works that have followed unconventional paths, but are not necessarily of lesser value (Fig. 2).

In the following sections of the methodological section, we will describe the tools that will allow us to construct our narratives, focusing both on indicators and information sources. Although there is a wide range of both, it is crucial to recognize that not all indicators are meaningful and not all sources are suitable for constructing narratives aligned with the recommendations of Table 5. Our intention is not to provide an exhaustive list, which would be materially impossible, but to select the most representative indicators and sources and explore how we can use them effectively.

5.1. Bibliometric indicators

A common argument in manifestos like those of Leiden and DORA is the problem of basing decisions solely on one indicator. Against this, the use of multiple indicators that provide a broader and more detailed view of what is being evaluated is recommended. Indeed, this claim is not new. It is necessary to bear in mind that, far from seeking a perfect indicator adaptable to any scenario, it is crucial to properly understand the different indicators, especially their limitations. The key lies in integrating them, overcoming their potential deficiencies, rather than discarding them because of their limitations (Martin, 1996). This integration allows for a balanced evaluation that more accurately reflects the complexity of the phenomena evaluated. Therefore, the focus should be on how various indicators can complement each other to provide a more complete and holistic view, which in turn leads to more informed and effective decisions.

From the outset in the literature, there have been proposals on how to classify bibliometric indicators into different dimensions (production, collaboration, etc.) and use them together. One of the earliest mentions of the multidimensionality of indicators can be found in Moravcsik (1984), and since then, there have been numerous proposals for classifying citation indicators and the concurrent use of various indicators (Vinkler, 1988; Moed et al., 1995). With the advent of digital metrics (Bollen et al., 2009) and altmetrics (Priem et al., 2011), these taxonomies, as well as the number of indicators and dimensions, have substantially expanded. It is now possible to quantify not only the scientific impact but also the usage of the document in other contexts (Sugimoto et al., 2017). In this regard, Moed (2017b) provides an integrative example of facet-based indicators, offering metrics that reflect the following dimensions: scientific-scholarly, educational/teaching, social, technological, economic, and cultural. Considering the tradition of these contributions, we have established four basic dimensions for the different indicators that will be used in our proposal for Narrative Bibliometrics.

- **Scientific impact:** This dimension focuses on measuring the relevance and influence that a publication has within the scientific community. It uses metrics based on the number of times the work is cited in other academic works, providing a direct reflection of both the attention and usage it is receiving from the community.
- **Relevance of the medium:** This dimension pays attention to the environment in which the work is published, considering both quantitative and qualitative aspects. Quantitative criteria derive from the scientific impact of its publications, while qualitative criteria address editorial policies and the quality of the medium, reflecting the importance of context in the valuation of the publication.
- **Social attention:** This dimension addresses the dissemination of the publication in social media based on mentions it receives on digital platforms and conversations on social networks, thereby reflecting attention from agents who may be outside academia.

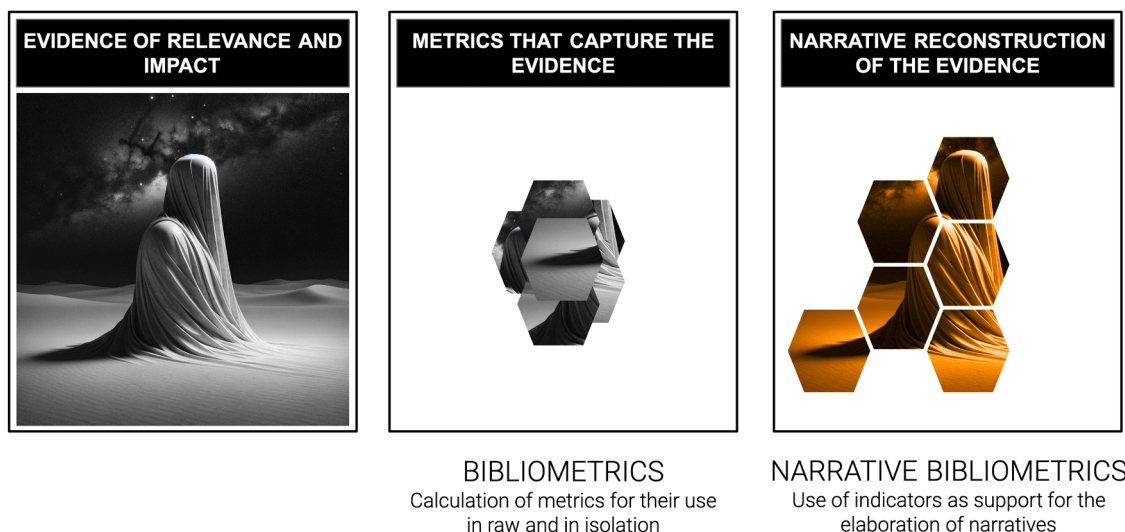


Fig. 2. Narrative Bibliometrics as a process of reconstructing the evidence.

- **Usage and visibility:** This dimension concentrates on the accessibility and reach of the publication in the digital world through metrics that directly address the usage of the work and the traffic generated by it. This provides a reflection of the access and reach of the publication.

The selection of indicators in Table 6 consolidates the main quantitative metrics to support Narrative Bibliometrics, covering the four dimensions. Following DORA's recommendations, these dimensions should be combined and used together in evaluations, being particularly cautious with relevance indicators for the medium, whose role must be minimized (Table 4) and necessarily complemented by other indicators. In addition to integrating indicators from various dimensions into a narrative and providing coherence, it is necessary to explore and dissect the values behind them to better understand the scope of the publication and provide a broader context to the metrics from which they were derived. This method is particularly relevant for deeply understanding the audiences and agents paying attention to our work.

It is also fundamental to recognize that all indicators are constructions that closely depend on specific contexts, that is, on the sources in which they are elaborated. With the rise of scientometric information sources, ranging from databases and search engines to altmetric data aggregators, the options have multiplied and diversified. A clear example of this evolution is that, before the end of the 20th century, Web of Science (then known as Web of Knowledge) was practically the only citation database. However, since then, we have witnessed the emergence of dozens of databases (Gusenbauer, 2022) with OpenAlex as the most recent and relevant novelty (Priem et al., 2022). This has resulted in a diversity of indicators, common and specific to each source, all dependent on the sources on which they are calculated. The inherent advantages and limitations of these tools are transferred to the indicators they produce. Although such proliferation is positive, as it offers customized and accessible solutions for users, it also constitutes significant risks for Bibliometrics, and particularly Narrative Bibliometrics.

5.2. The sources

As varied as the indicators are, so too are the sources from which they derive. In addition to the clear explosion in bibliographic database sources, the main source of information for this type of data, there has also been a diversification in bibliometric data products. With this, we can establish the following classification of sources for obtaining metrics (Table 7 and Table 8).

Table 6
Main/selected indicators and useful sources for their use.

<i>Dimension</i>	<i>Indicators</i>	<i>Sources</i>	<i>Type</i>	<i>Interpretation</i>
SCIENTIFIC IMPACT	Total citations	WoS / Scopus / Dimensions / Dialnet Métricas / GS	Raw indicator	Total number of times a publication is referenced by others
	Total citations (excluding self-citations)	WoS / Scopus / Dialnet Métricas	Raw indicator	Total number of times a publication is referenced by others, excluding any by the authors themselves
	Normalized citations	WoS / Scopus / Dimensions / Dialnet Métricas	Normalized value	Difference between the number of citations and the global average for publications with similar characteristics
	Citation percentile	InCites / Scopus / Dialnet Métricas	Normalized value	Position of the publication in terms of citation numbers compared to others in its field
RELEVANCE OF THE MEDIUM	Reviews	Dialnet	Raw indicator	Total number of times a book is reviewed
	Journal impact	JCR / SJR / Scopus / IDR	Raw indicator	Impact measure based on the citations received by the publications within a time window
	Journal percentile	JCR / Scopus / IDR	Normalized value	Position of the journal in terms of citation numbers compared to others in its field
	Journal quartile	JCR / SJR / Scopus / IDR	Normalized value	Quartile position of the journal in the impact ranking of its field
SOCIAL ATTENTION	Journal ranking	JCR / SJR / Scopus / IDR	Raw indicator	Position of the journal in the impact ranking of its field
	Social media mentions	Altmetric.com / PlumX	Raw indicator	Total number of times a publication is mentioned on a social network
	News mentions	Altmetric.com / PlumX	Raw indicator	Total number of times a publication is referenced in digital press
	Mentions in policy reports	Altmetric.com / PlumX	Raw indicator	Total number of times a publication is referenced in public policy reports
USAGE AND VISIBILITY	Wikipedia mentions	Altmetric.com / PlumX	Raw indicator	Total number of times a publication is referenced in Wikipedia articles
	Mendeley readers	Mendeley / Altmetric.com / PlumX	Raw indicator	Total number of Mendeley users who have added the publication to their library
	Academic links	Altmetric.com / PlumX	Raw indicator	Total number of times a publication is referenced on curated academic websites
	Downloads	Web analytics tools	Raw indicator	Total number of times a publication has been downloaded
	Presence in libraries	REBIUN / WorldCat	Raw indicator	Total number of libraries that have the book in their catalog

Table 7
Classification of bibliometric information sources.

Bibliographic reference and citation databases	Search engines
Traditional platforms specialized in the collection and organization of bibliographic references of scientific publications and their citations. <i>For example, Web of Science and Dialnet.</i>	Platforms that allow for searching academic and scientific literature across the Internet. <i>For example, Google Scholar and PubMed.</i>
Data aggregators	Indexes
Specialized platforms that compile and integrate data related to scientific publications. <i>For example, Altmetric.com and PlumX.</i>	Tools that provide curated lists of journals and books including detailed information. <i>For example, Journal Citation Reports and DOAB.</i>
Combined sources	Open Access Repositories
These platforms combine more than one source at the same time, especially bibliographic databases and data aggregators. <i>For example, Dimensions+Altmetric.com and Scopus+PlumX.</i>	Digital platforms that store and allow access to academic works in open access, including thematic repositories specialized in specific areas and institutional repositories. <i>For example, Zenodo.</i>
Bibliometric suites	Complementary sources
Highly specialized tools focused exclusively on bibliometric and scientometric analysis. <i>For example, InCites and SciVal.</i>	A variety of resources not focused on academic literature but through which it interacts. <i>For example, WorldCat and Wikipedia.</i>

6. From metrics to narratives

In developing these narratives, the primary risk is turning the diversity of metrics and sources into a marketplace rush for the best offer to craft persuasive narratives. Narrative Bibliometrics should always be a process that moves from evidence to indicators, not vice versa. It must be handled with caution to avoid mixing elements that, far from complementing each other, deepen the same message. Here lie two very common risks: redundancy and incoherence (Fig. 3).

The first has to do with the obsession with metrics (Muller, 2019), leading to their excessive use and lack of true relevance, often reiterating the same message. The second is more severe, as it involves constructing a Frankenstein with completely decontextualized indicators that cannot be jointly interpreted. This phenomenon occurs when metrics that only make sense within their original context or specific database are mixed as part of a whole. The key, therefore, is to use metrics in an informed and contextual manner, recognizing their limitations and avoiding falling into the trap of overinterpretation, underinterpretation, or misinterpretation.

Once we have the necessary metrics, it is time to exploit the evidence reflected by the metrics by constructing narratives. Table SM1, included in *Supplementary Material* (doi:10.5281/zenodo.11127617), presents various examples showing how exhaustive narratives can be established from specific indicators collected in Table 6, contextualizing such indicators and highlighting their value. Table SM1 provides practical examples that can guide the application of narrative metrics, facilitating a comprehensive assessment of a study's impact and relevance across different dimensions of impact. The narratives associated with each indicator enhance visibility into the diverse aspects of scientific research and its implications, promoting a richer understanding among scholars, policymakers, and the academic community. Narrative bibliometrics also offers us the possibility to explain and justify the decisions we have to made for the calculation of our indicators, and how them can affect the retrieved values. Moreover, a science like bibliometrics that is mainly based on statistics should always indicate the validity of its values, both of its figures as well as of its decimals, if they have any significance.

For example, it would be very necessary to “narrate” how the values obtained for the calculation of the normalized citation indicators, such as the CNCI or the percentiles, would be affected by the decision we have taken by 1) selecting a particular classification scheme, 2) selecting one classification schema based on journal level or on document level; and 3) using the selected classification schema on a particular aggregation level (macro, meso or micro). Actually, for the bibliometrician community, it is crucial to be able to estimate the effect that these decisions are expected to have on the values calculated for their impact measures. The errors that can be made in the calculations of our bibliometric indicators when making decisions about the type of documents to be analysed or the type of classification chosen have been recently analysed and discussed (see Robinson et al., 2023). The authors conclude that their findings underscore the importance of responsible metric use in research evaluation, providing valuable insights for both bibliometricians and consumers of such data. Narrative bibliometrics would be the ideal instrument to undertake such an arduous but necessary task.

7. Conclusions

This paper synthesises key developments in scholarly evaluation, drawing on influential manifestos such as DORA and CoARA, alongside contemporary scientific policies. It explores the evolution of academic assessment, highlighting the transition towards more ethical and responsible evaluation practices, notably within European scientific frameworks. The introduction of Narrative Bibliometrics—a paradigm that integrates bibliometric indicators with academic narratives—further enriches this transition. Developed by established researchers and practitioners from European universities, this innovative approach emphasises the application of metrics while respecting the contextual nuances of scientific contributions. Its recent adoption and seamless integration into the evaluation processes carried out by the Spanish evaluation agency (ANECA) effectively demonstrate the proposal's efficacy in enhancing professional research evaluation.

The Narrative Bibliometrics operational framework defines core principles such as applicability, uniqueness, adaptability, comparability, and contextuality. These principles are essential for interpreting indicators within diverse academic landscapes. By

Table 8

Main information sources and their bibliographic and metric characteristics.

SOURCE	TYPE	BIBLIOGRAPHIC CONTENT				METRICS		
		DOCUMENTARY SPECIALIZATION	THEMATIC SPECIALTY	LINGUISTIC BIAS	METADATA CURATION	METRICS LEVEL	IMPACT TYPE	NORMALIZED METRICS
 Web of Science [®]	Bibliographic database	Articles	Natural and Physical Sci. Health and Technology	English	High	Document	Scientific	No
 Dialnet	Bibliographic database	Articles Books	Social Sciences Humanities	Spanish	Normal	Document Medium	Scientific	No
 Scopus [®] 	Combined source	Articles	Natural and Physical Sci. Health and Technology	English	High	Document Medium	Scientific Social	Yes
 Dimensions Altmetric	Combined source	Articles	Natural and Physical Sci. Health and Technology	English	Normal	Document	Scientific Social	Yes
 Clarivate InCites [™]	Bibliometric suite	Articles	Natural and Physical Sci. Health and Technology	English	High	Document Medium	Scientific	Yes
 Dialnet métricas	Bibliometric suite	Articles Books	Social Sciences Humanities	Spanish	Normal	Document Medium	Scientific	Yes
 Clarivate Analytics Journal Citation Reports	Index	Journals	Natural and Physical Sci. Health and Technology	English	High	Medium	Scientific	Yes
 OCLC WorldCat [®]	Complementary source	Books	Social Sciences Humanities	Multilingual	Normal	Document	Social	No

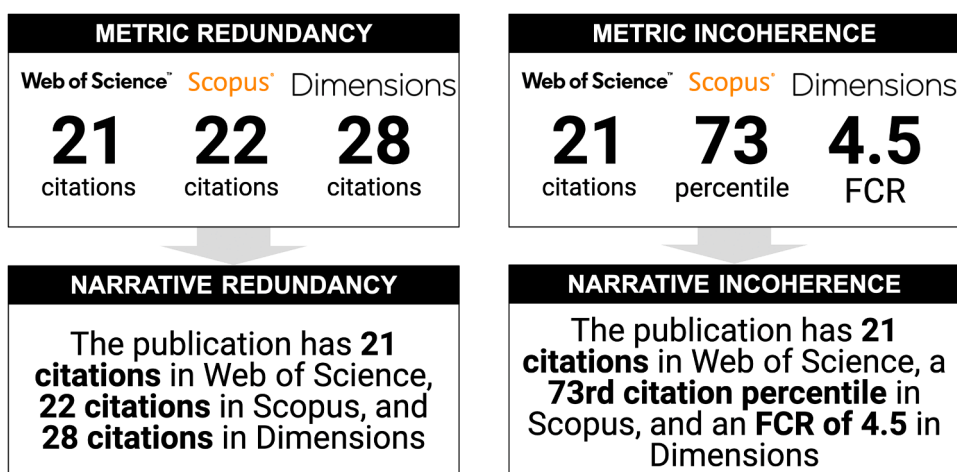


Fig. 3. Main problems of redundancy and incoherence in narrative bibliometrics.

categorising indicators into four dimensions—scientific impact, publication channel relevance, social attention, and social interaction—the framework aligns with the enhanced research assessment methods implemented by ANECA. Moreover, this approach is strategically designed to complement, rather than replace, reasoned review reports. It underscores the importance of bibliometric indicators in areas critical for demonstrating the impact, relevance, and influence of scientific contributions. Ultimately, this study not only promotes the responsible use of bibliometric indicators but also serves as both a practical tool for evaluators and a conceptual framework for the broader bibliometrics community, thereby facilitating the advancement of scholarly assessment practices.

Supplementary material

Table SM1, providing examples of narrative construction from specific metrics, can be consulted on Zenodo at the doi:[10.5281/zenodo.11127617](https://doi.org/10.5281/zenodo.11127617).

CRediT authorship contribution statement

Daniel Torres-Salinas: Writing – original draft, Supervision, Project administration, Methodology, Investigation, Formal analysis. **Enrique Orduña-Malea:** Writing – original draft, Validation, Conceptualization. **Ángel Delgado-Vázquez:** Writing – original draft, Validation, Conceptualization. **Juan Gorraiz:** Writing – review & editing, Validation. **Wenceslao Arroyo-Machado:** Writing – review & editing, Writing – original draft, Visualization, Formal analysis.

Declaration of competing interest

The author is an Editorial Board Member/Editor-in-Chief/Associate Editor/Guest Editor for *[Journal of Informetrics]* and was not involved in the editorial review or the decision to publish this article.

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:

Daniel Torres-Salinas, Enrique Orduña-Malea, and Ángel Delgado-Vázquez belong to the Commission for Evaluation and Monitoring of the Spanish accreditation system at the National Agency for Quality Assessment and Accreditation (ANECA).

References

- Abadal, E. (2021). Ciencia abierta: Un modelo con piezas por encajar. *Arbor*, 197(799). <https://doi.org/10.3989/arbor.2021.799003>. Article 799.
- Aguillo, I. F. (2022). Bibliometría sofisticada. *Anuario ThinkEPI*, 16. <https://doi.org/10.3145/thinkepi.2022.e16e28>
- Almind, T. C., & Ingwersen, P. (1997). Informetric analyses on the world wide web: Methodological approaches to «webometrics». *Journal of Documentation*, 53(4), 404–426. <https://doi.org/10.1108/EUM00000000007205>
- Bollen, J., Van de Sompel, H., Smith, J. A., & Luce, R. (2005). Toward alternative metrics of journal impact: A comparison of download and citation data. *Information Processing & Management*, 41(6), 1419–1440. <https://doi.org/10.1016/j.ipm.2005.03.024>
- Bollen, J., Sompel, H. V. de, Hagberg, A., & Chute, R. (2009). A principal component analysis of 39 scientific impact measures. *PloS one*, 4(6), e6022. <https://doi.org/10.1371/journal.pone.0006022>
- Bornmann, L., & Williams, R. (2017). Can the journal impact factor be used as a criterion for the selection of junior researchers? A large-scale empirical study based on ResearcherID data. *Journal of Informetrics*, 11(3), 788–799. <https://doi.org/10.1016/j.joi.2017.06.001>
- Byrne, D. (2023). Could new 'narrative' CVs transform research culture? *Nature*. <https://doi.org/10.1038/d41586-023-03219-8>
- Cobey, K. (2024). Impact factors are outdated, but new research assessments still fail scientists. *Nature*, 628(8006). <https://doi.org/10.1038/d41586-024-00899-8>, 9–9.

- Daraio, C., & Glänzel, W. (Eds.). (2020). *Evaluative informetrics: The art of metrics-based research assessment: Festschrift in honour of Henk F. Moed*. Springer International Publishing. <https://doi.org/10.1007/978-3-030-47665-6>.
- Fritch, R., Hatch, A., Hazlett, H., & Vinkenburg, C. (2021). *Using narrative CVs: Process optimization and bias mitigation*. Zenodo. <https://doi.org/10.5281/zenodo.5799414>
- Garfield, E. (2002). *The Impact Factor and Using It Correctly*. <https://www.semanticscholar.org/paper/The-Impact-Factor-and-Using-It-Correctly-Garfield/39e8b0e8c870b9cb43b3687d1689f27d73722f17>.
- Garfield, E. (2006). The history and meaning of the journal impact factor. *JAMA-Journal of The American Medical Association*, 295(1), 90–93. <https://doi.org/10.1001/jama.295.1.90>
- Gusenbauer, M. (2022). Search where you will find most: Comparing the disciplinary coverage of 56 bibliographic databases. *Scientometrics*, 127(5), 2683–2745. <https://doi.org/10.1007/s11192-022-04289-7>
- Hazlett, H. (2022). Changing the narrative: Considering common principles for the use of narrative CVs in grant evaluation. *DORA*. <https://sfedora.org/2022/06/06/changing-the-narrative-considering-common-principles-for-the-use-of-narrative-cvs-in-grant-evaluation/>.
- Hicks, D., Wouters, P., Waltman, L., de Rijcke, S., & Rafols, I. (2015). Bibliometrics: The Leiden manifesto for research metrics. *Nature*, 520(7548). <https://doi.org/10.1038/520429a>. Article 7548.
- Larivière, V., & Sugimoto, C. R. (2019). The journal impact factor: a brief history, critique, and discussion of adverse effects. In W. En, H. F. Glänzel, U. Schmoch Moed, & M. Thelwall (Eds.), *Springer handbook of science and technology indicators* (pp. 3–24). Springer International Publishing. https://doi.org/10.1007/978-3-030-02511-3_1.
- Lee, C. J., Sugimoto, C. R., Zhang, G., & Cronin, B. (2013). Bias in peer review. *Journal of the American Society for Information Science and Technology*, 64(1), 2–17. <https://doi.org/10.1002/asi.22784>
- Leydesdorff, L. (2012). Alternatives to the journal impact factor: I3 and the top-10% (or top-25%) of the most-highly cited papers. *Scientometrics*, 92(2), 355–365. <https://doi.org/10.1007/s11192-012-0660-6>
- Martin, B. R. (1996). The use of multiple indicators in the assessment of basic research. *Scientometrics*, 36(3), 343–362. <https://doi.org/10.1007/BF02129599>
- Moed, H., Debruin, R., & Van Leeuwen, T. (1995). New bibliometric tools for the assessment of national research performance—Database description, overview of indicators and first applications. *Scientometrics*, 33(3), 381–422. <https://doi.org/10.1007/BF02017338>
- Moed, H. F. (2010). Measuring contextual citation impact of scientific journals. *Journal of Informetrics*, 4(3), 265–277. <https://doi.org/10.1016/j.joi.2010.01.002>
- Moed, H. F. (2017a). *Applied evaluative informetrics* (1st ed.). Springer International Publishing. <https://doi.org/10.1007/978-3-319-60522-7>, 2017Imprint: Springer.
- Moed, H. F. (2017b). Multi-dimensional Research Performance. In En H. F. Moed (Ed.), *Applied evaluative informetrics* (pp. 45–59). Springer International Publishing. https://doi.org/10.1007/978-3-319-60522-7_3.
- Moed, H.F. (2020). *Appropriate Use of Metrics in Research Assessment of Autonomous Academic Institutions* (1). 2(1), Article 1. <https://doi.org/10.29024/sar.8>.
- Mongeon, P., & Paul-Hus, A. (2016). The journal coverage of Web of Science and Scopus: A comparative analysis. *Scientometrics*, 106(1), 213–228. <https://doi.org/10.1007/s11192-015-1765-5>
- Moravcsik, M. J. (1984). Life in a multidimensional world. *Scientometrics*, 6(2), 75–85. <https://doi.org/10.1007/BF02021280>
- Muller, J. Z. (2019). *The tyranny of metrics*. Princeton University Press (First paperback edition).
- Narin, F. (1976). *Evaluative bibliometrics: The use of publication and citation analysis in the evaluation of scientific activity*. Computer Horizons.
- Orduña-Malea, E., & Bautista-Puig, N. (2024). Research assessment under debate: Disentangling the interest around the DORA declaration on Twitter. *Scientometrics*, 129(1), 537–559. <https://doi.org/10.1007/s11192-023-04872-6>
- Pérez Esparrells, C., Bautista-Puig, N., & Orduña-Malea, E. (2022). *La evaluación pública de la investigación científica en el contexto internacional: Posibilidades y límites. Informe I* [Report]. <https://repositorio.uam.es/handle/10486/705530>.
- Petersohn, S., & Heinze, T. (2018). Professionalization of bibliometric research assessment. Insights from the history of the Leiden Centre for Science and Technology Studies (CWTS). *Science and Public Policy*, 45(4), 565–578. <https://doi.org/10.1093/scipol/scx084>
- Priem, J., & Hemminger, B. H. (2010). Scientometrics 2.0: New metrics of scholarly impact on the social Web. *First Monday*. <https://doi.org/10.5210/fm.v15i7.2874>
- Priem, J., Taraborelli, D., Groth, P., & Neylon, C. (2011). *Altmetrics: A manifesto*. DigitalCommons@University of Nebraska - Lincoln. <https://digitalcommons.unl.edu/scholcom/185>.
- Priem, J., Piwowar, H., & Orr, R. (2022). *OpenAlex: A fully-open index of scholarly works, authors, venues, institutions, and concepts* (arXiv:2205.01833). arXiv. <http://arxiv.org/abs/2205.01833>.
- Rafols, I., & Stirling, A. (2021). Designing indicators for opening up evaluation: Insights from research assessment. En *A research agenda for evaluation*, 165–188 <https://www.elgaronline.com/configurable/content/edcoll002f97818391010760002f9781839101076.00015.xml?tc=edcoll%24002f9781839101076%24002f9781839101076.00015.xml>
- Robinson-García, N., Torres-Salinas, D., Vargas-Quesada, B., Chinchilla-Rodríguez, Z., & Gorraiz, J. (2023). Errors of measurement in scientometrics: Identification and calculation of systematic errors. In *Proceedings of ISSI 2023 – the 19th international conference of the international society for scientometrics and informetrics* (pp. 387–393). <https://doi.org/10.5281/zenodo.8428899>. Vol. 2Presented at the ISSI 2023.
- Singh Chawla, D. (2022). Swiss funder unveils new CV format to make grant evaluation fairer. *Nature*, 606(7916), 1033–1034. <https://doi.org/10.1038/d41586-022-01599-x>
- Sugimoto, C. R., Work, S., Larivière, V., & Haustein, S. (2017). Scholarly use of social media and altmetrics: A review of the literature. *Journal of the Association for Information Science and Technology*, 68(9), 2037–2062. <https://doi.org/10.1002/asi.23833>
- Thelwall, M. (2008). Bibliometrics to webometrics. *Journal of Information Science*, 34(4), 605–662.
- Thelwall, M., & Fairclough, R. (2015). Geometric journal impact factors correcting for individual highly cited articles. *Journal of Informetrics*, 9(2), 263–272. <https://doi.org/10.1016/j.joi.2015.02.004>
- Torres Salinas, D. (2024). *Principios de bibliometría evaluativa*. EPI-UOC.
- Torres-Salinas, D. (2023a). Entre métricas y narraciones: Definición y aplicaciones de la Bibliometría Narrativa. *Anuario ThinkEPI*, 17. <https://doi.org/10.3145/thinkepi.2023.e17a30>
- Torres-Salinas, D., Arroyo-Machado, W., & Robinson-García, N. (2023b). Bibliometric denialism. *Scientometrics*, 128(9), 5357–5359. <https://doi.org/10.1007/s11192-023-04787-2>
- Torres-Salinas, D., Cabezas-Clavijo, Á., & Jiménez-Contreras, E. (2013). Altmetrics: New indicators for scientific communication in Web 2.0. *Comunicar*, 21(41), 53–60. <https://doi.org/10.3916/C41-2013-05>
- Torres-Salinas, D., Robinson-García, N., Herrera-Viedma, E., & Jiménez-Contreras, E. (2018). Consideraciones metodológicas sobre uso del impacto normalizado en convocatorias Severo Ochoa y María de Maetzu. In *Profesional de la Información*, 27. <https://doi.org/10.3145/epl.2018.mar.15>. Article 2.
- UKRI. (2023). *Résumé for Research and Innovation (R4RI): Guidance*. <https://www.ukri.org/apply-for-funding/how-to-apply/resume-for-research-and-innovation-r4ri-guidance/>.
- Van Raan, A. (2019). Measuring science: Basic principles and application of advanced bibliometrics. In W. En, H. F. Glänzel, U. Schmoch Moed, & M. Thelwall (Eds.), *Springer handbook of science and technology indicators* (pp. 237–280). Springer.
- Vinkler, P. (1988). An attempt of surveying and classifying bibliometric indicators for scientometric purposes. *Scientometrics*, 13(5–6), 239–259. <https://doi.org/10.1007/BF02019961>
- Wilsdon, J., Allen, L., Belfiore, E., Campbell, P., Curry, S., Hill, S. et al. (2015). *The metric tide: Report of the independent review of the role of metrics in research assessment and management*. <https://doi.org/10.13140/RG.2.1.4929.1363>.
- Woolston, C. (2022). Time to rethink the scientific CV. *Nature*, 604(7904), 203–205. <https://doi.org/10.1038/d41586-022-00928-4>